

COMMENT J'EXPLORE...

Les critères de jugement dans les essais cliniques : réflexions à propos d'études récentes de prévention cardio-vasculaire

A.J. SCHEEN (1)

RÉSUMÉ : La comparaison de deux interventions thérapeutiques nécessite le recours à des critères de jugement dans des essais cliniques randomisés, ce qui constitue la base même de la médecine factuelle. Il est convenu de privilégier les critères forts, comme les événements cliniques majeurs et la mortalité plutôt que les critères dits accessoires. Dans les grands essais de prévention cardio-vasculaire, il est devenu habituel d'utiliser des critères de jugement primaire(s) et secondaire(s) et de recourir à des critères cliniques composites. Ces derniers ont l'avantage d'augmenter le nombre d'événements, mais l'inconvénient de diluer l'effet thérapeutique et d'exposer à certains biais. Nous discutons cette problématique importante à la lumière de la publication des résultats controversés de quelques grands essais de prévention cardio-vasculaire récents.

MOTS-CLÉS : *Critères de jugement – Efficacité – Essais cliniques contrôlés – Médecine factuelle – Médicament – Prévention cardio-vasculaire*

HOW I EXPLORE... OUTCOMES IN CLINICAL TRIALS : COMMENTS
ABOUT RECENT STUDIES OF CARDIOVASCULAR PREVENTION.

SUMMARY : The comparison of two therapeutic interventions requires the use of endpoints in randomized clinical trials, which supports evidence-based medicine. It is preferable to consider hard endpoints, such as major clinical events and mortality rather than biological or surrogate endpoints. In large cardiovascular prevention trials, it is now well accepted to analyze primary and secondary endpoints as well as to use combined clinical endpoints. Such composite endpoints offer the advantage of increasing the number of events, but have the disadvantage of diluting the therapeutic effect and of exposing to possible bias. We discuss this crucial issue at the light of controversial results from a few recent important cardiovascular prevention trials.

KEYWORDS : *Cardiovascular prevention – Drug – Evidence-based medicine – Outcomes – Randomised clinical trials – Efficacy*

INTRODUCTION

Les essais cliniques contrôlés sont le fondement de la médecine factuelle ou «Evidence-Based-Medicine» (EBM) (1). Ceux-ci visent à démontrer la supériorité d'une approche thérapeutique par rapport à un comparateur, qu'il soit un placebo ou un traitement de référence. La construction de ces essais cliniques se fonde sur la méthode «PICO» : P pour «Patient», I pour «Intervention», C pour Comparateur (ou Contrôle) et O pour « Outcome » (2, 3). L'«Outcome», c'est-à-dire le critère de jugement utilisé, est évidemment crucial pour démontrer la supériorité d'un traitement par rapport à un autre (4). Les essais cliniques dans le domaine de la prévention cardio-vasculaire sont particulièrement nombreux. Alors que, pendant longtemps, seuls des critères de jugement accessoires («soft endpoints»), à savoir la diminution des facteurs de risque (baisse de la cholestérolémie, de la pression artérielle, de la glycémie) étaient utilisés, le recours à des critères de jugement cliniques forts («hard endpoints»), à savoir la diminution de l'incidence des événements cardio-vasculaires et de la mortalité cardio-vasculaire ou globale, est maintenant privilégiée. Ces critères ont, bien évidemment, un poids plus important dans l'opinion finale que se font les commissions d'enregistrement et de remboursement des médicaments, d'une part, les cliniciens prescripteurs, d'autre part (5).

Les résultats de quelques grands essais cliniques de prévention cardio-vasculaire ont été rapportés en 2005 en ciblant soit l'hypercholestérolémie (TNT, IDEAL, FIELD) (6-8), soit l'hypertension artérielle (ASCOT) (9), soit le diabète (PROactive) (10). Ces études, toutes remarquables, ont déjà fait l'objet d'une présentation détaillée dans la revue (11-15). Dans la plupart de ces essais contrôlés, le critère de jugement principal, par ailleurs composite, n'était pas significativement différent entre les deux interventions thérapeutiques alors que certains critères de jugement secondaire, composites ou non, l'étaient. Ces discordances ont suscité et suscitent encore une vive controverse. Aussi, il nous a paru intéressant de partager quelques réflexions sur la hiérarchisation des critères de jugement dans les essais cliniques contrôlés (16) ainsi que sur les avantages et les limites des critères de jugement composite (17).

TYPES DE CRITÈRES DE JUGEMENT

Les critères de jugement peuvent être très variables, allant de simples critères biologiques à des événements cliniques majeurs dont les décès, en passant par des critères d'évaluation intermédiaire faisant appel à une imagerie médicale de plus en plus performante.

RÉDUCTION DU FACTEUR DE RISQUE DE LA MALADIE

Les premiers essais ont analysé les effets des médicaments sur les facteurs de risque de la maladie cardio-vasculaire, essentiellement l'élévation de la pression artérielle et l'hypercholes-

(1) Professeur Ordinaire, Université de Liège, Chef de Service, Service de Diabétologie, Nutrition et Maladies métaboliques, Unité de Pharmacologie clinique, Département de Médecine.

térolémie, plus rarement l'hyperglycémie. L'avantage de cette stratégie est que la variable testée est continue, c'est-à-dire que chaque sujet présente une modification qui peut être analysée de façon appropriée sur le plan statistique (4). Cette approche est justifiée par la relation forte, mise en évidence dans de nombreuses études épidémiologiques, entre le risque d'événement cardio-vasculaire et le niveau du facteur de risque. Néanmoins, il est de plus en plus admis qu'il convient de privilégier des essais qui, non seulement analysent des modifications de facteurs de risque, mais aussi qui évaluent les effets sur l'expression clinique de la maladie elle-même (5).

AMÉLIORATION D'UN MARQUEUR INTERMÉDIAIRE DE LA MALADIE

Le recours à un marqueur intermédiaire de la maladie est justifié par l'objectif de démontrer un effet du médicament testé sur l'évolution même de la maladie (plutôt que sur un simple facteur de risque) et ce, dans un laps de temps raisonnable (16). Dans le domaine cardio-vasculaire, le marqueur intermédiaire reconnu par la «Food and Drug Administration» (FDA) est le rapport intima/media évalué par échographie au niveau de l'artère carotide. Un autre marqueur potentiel, pris en considération dans certaines études cardiologiques, est le degré de sténose coronarienne évalué par artériographie. Ce marqueur est, à présent, quasi abandonné au profit d'une mesure directe de l'athérome coronarien grâce à la technique d'échographie endovasculaire (IVUS) (18). Ce type de marqueur intermédiaire offre l'avantage de pouvoir être apprécié après un suivi de moindre durée que les événements cliniques et de pouvoir être analysé de façon quantitative chez tous les patients inclus dans l'étude. Il faut cependant bien reconnaître qu'il s'agit là de critères intermédiaires («surrogate endpoints»). On comprendra que la réduction de l'incidence des événements cliniques doit être considérée comme l'objectif ultime en terme de prévention cardio-vasculaire (19).

DIMINUTION DES ÉVÉNEMENTS CLINIQUES LIÉS À LA MALADIE

Les événements cliniques sont certes les plus importants à prendre en considération, mais ils ne surviennent pas chez tous les patients et, de plus, ils peuvent n'apparaître qu'après un délai relativement important. Pour avoir une puissance statistique suffisante, les essais visant à analyser ce critère fort doivent inclure un grand nombre de sujets suivis pendant une durée suffisamment longue. C'est d'autant plus le cas que

les sujets inclus sont à relativement faible risque (par exemple, sujets jeunes, de sexe féminin et/ou en prévention primaire). En outre, les événements cliniques pris en considération peuvent être très différents puisqu'il peut s'agir de décès (coronariens, cardio-vasculaires, ou totaux), d'événements cliniques majeurs (infarctus non mortels, accidents vasculaires cérébraux), d'accidents moins sévères (angor instable ou non, décompensation cardiaque) ou encore de procédures de revascularisation (coronaire, carotidienne ou périphérique). Il est évident que tous ces événements n'ont pas le même poids, ni pour le patient, ni en terme de santé publique, ni enfin sur le plan des répercussions économiques. Par ailleurs, leur physiopathologie peut être très différente comme, par exemple, celle de l'insuffisance coronaire et celle de la décompensation cardiaque; dès lors, il est vraisemblable que le médicament testé, en fonction de son mécanisme d'action, n'ait pas nécessairement le même impact selon le critère pris en compte. Ainsi, dans l'étude VALUE comparant le valsartan, un antagoniste des récepteurs AT1 de l'angiotensine, et l'amlodipine, une dihydropyridine bloquant les canaux calciques, chez des sujets hypertendus, il n'y avait pas de différence significative entre les deux groupes traités pour ce qui concerne le critère d'évaluation composite cardio-vasculaire primaire (20); l'analyse des différentes composantes indiquait, néanmoins, un moindre risque de décompensation cardiaque sous valsartan, mais une diminution significative de l'incidence des infarctus du myocarde sous amlodipine, comme attendu au vu des mécanismes d'action spécifiques des deux molécules testées. Enfin, il apparaît que certains des critères utilisés sont exclusivement dépendant de l'évolution de la maladie, alors que d'autres (comme les procédures de revascularisation) sont, au moins partiellement, influencés par l'attitude du médecin en charge du malade («opérateur-dépendants»).

HIÉRARCHIE DES CRITÈRES DE JUGEMENT

Une étude clinique, si elle est suffisamment ambitieuse et si elle a une envergure suffisante, essaie de répondre à plusieurs questions. Il est d'usage de hiérarchiser les critères de jugement en fonction de leur importance présumée : il est alors recommandé de les répartir en critère de jugement principal («primary endpoint») et en critères de jugement secondaires (voire même tertiaires).

CRITÈRE DE JUGEMENT PRINCIPAL

Le critère de jugement principal devrait être celui qui est le plus important du point de vue clinique ou en terme de santé publique. Force est cependant de reconnaître que le critère choisi peut dépendre d'autres facteurs, importants pour la firme pharmaceutique qui supporte l'essai. Les deux facteurs habituellement pris en considération sont, d'une part, l'importance supputée de la différence de survenue du critère entre les deux conditions expérimentales (plus la différence potentielle est grande, plus elle sera facile à mettre en évidence), d'autre part, la probabilité d'avoir une puissance statistique suffisante. Cette dernière dépendra directement du nombre d'événements susceptibles de survenir en cours d'étude. Pour augmenter ce paramètre, il conviendra, si l'on veut limiter le nombre de patients à inclure et la durée du suivi à imposer, de recourir à un critère de jugement composite (combinant plusieurs types d'événements : voir ci-dessous).

Il est intéressant de noter que les critères combinés de jugement principal utilisés dans les différentes études cliniques récentes de prévention cardio-vasculaire varient selon les études considérées (Tableau I). Fait encore plus étonnant, le critère de jugement principal est susceptible d'être modifié en cours même d'étude, comme cela s'est avéré le cas dans certaines études récentes. Il est cependant essentiel d'effectuer cette modification de façon explicite, le plus tôt possible et, en tous cas, avant l'ouverture du code de randomisation et en prévenant les instances officielles *in tempore non suspecto*. Ainsi, le critère d'évaluation principal a été modifié dans le cadre de l'étude FIELD : le critère proposé initialement était la mortalité coronarienne tandis que le critère principal amendé, publié bien avant la fin de l'étude (21), concernait à la fois les décès coronariens et les infarctus non mortels (8). Ce changement s'est révélé positif pour l'étude puisque, dans FIELD, l'incidence des décès coronariens n'a pas été réduite par le fénofibrate alors que l'incidence des

TABEAU I : COMPARAISON DES COMPOSANTES FORMANT LE CRITÈRE DE JUGEMENT COMPOSITE PRINCIPAL DANS QUELQUES GRANDS ESSAIS CONTRÔLÉS RÉCENTS DE PRÉVENTION CARDIO-VASCULAIRE CHEZ DES PERSONNES AVEC HYPERCHOLESTÉROLÉMIE (CHOL), HYPERTENSION ARTÉRIELLE (HTA) OU DIABÈTE DE TYPE 2 (DT2).

	TNT	IDEAL	FIELD	VALUE	ASCOT	PROACTIVE
Thématique	Chol	Chol	Chol	HTA	HTA	DT2
Comparateurs	A10 vs A80	S20-40 vs A80	Pbo vs Féno	Amlo vs Val	Amlo vs Atén	Pbo vs Pio
Patients (n)	10.001	8.888	9.795	15.245	19.257	5.238
Suivi (années)	4,9	4,8	5,0	4,2	5,5	2,8
Mortalité totale	-	-	-	-	-	+
Décès cardiaques	-	-	-	+	-	+
Décès coronariens	+	+	+	+	+	+
Infarctus non mortels	+	+	+	+	+	+
Infarctus silencieux	?	?	?	?	+	+
Mort subite avec ressuscitation	+	+	?	+	?	+
AVC	+	-	-	-	-	+
Angor instable (hospitalisation)	-	-	-	?	-	+
Insuffisance cardiaque (hospitalisation)	-	-	-	+	-	-
Toute revascularisation coronaire	-	-	-	-	-	+
Revascularisation coronaire en urgence	-	-	-	+	-	+
Revascularisation périphérique	-	-	-	-	-	+
Ischémie périphérique	-	-	-	-	-	+
Amputation au-dessus cheville	-	-	-	-	-	+
A10 : atorvastatine 10 mg Pbo : placebo A80 : atorvastatine 80 mg Féno : fénofibrate micronisé 200 mg Amlo : amlodipine 5-10 mg Val : valsartan 80-160 mg S20-40 : simvastatine 20-40 mg Atén : aténolol 50-100 mg Pio : pioglitazone 15-45 mg + : composante retenue dans le critère combiné - : composante non retenue ? : information non précisée						

infarctus non mortels l'a été de façon significative, mais, de façon insuffisante, pour que le critère de jugement principal atteigne le seuil de la significativité statistique entre le groupe fénofibrate, et le groupe placebo, chez les patients diabétiques de type 2 étudiés (8, 13). Dans l'étude PROactive, le critère de jugement principal n'a pas été modifié, mais un «critère de jugement secondaire principal» a été ajouté peu avant l'analyse finale des résultats (10). Ce critère, plus restrictif que le critère principal, comprend la mortalité toutes causes, l'infarctus du myocarde non fatal et l'AVC («stroke»), prenant donc en compte les événements cliniques majeurs et excluant les critères cliniques de diagnostic plus ambigu (angor instable) ou opérateur-dépendants (procédures de revascularisation ou d'amputation) (22). A nouveau, cette modification s'est avérée bénéfique puisque le critère d'évaluation primaire n'est pas statistiquement différent entre le groupe placebo et le groupe pioglitazone, alors que le critère secondaire principal a été positivement et significativement influencé par la pioglitazone chez les patients diabétiques de type 2 ayant participé à cette étude (10, 13). Ce double choix et la primauté apparente accordée par les investigateurs au critère secondaire principal par rapport au critère primaire ont fait l'objet d'une critique spécifique peu après la diffusion publique des résultats de PROACTIVE (23).

Pour montrer comment le choix du critère combiné principal peut être déterminant, il est intéressant de comparer les résultats des trois études récentes comparant directement deux traitements par statine chez des patients coronariens, à savoir TNT (6, 11), IDEAL (7, 12) et PROVE-IT (24) (cette dernière étude a été réalisée dans les suites d'un accident coronarien aigu). Ainsi, le critère principal pris en compte dans IDEAL ne montre pas de différence significative entre le groupe simvastatine 20-40 mg et le groupe atorvastatine 80 mg (HR = 0,89, $p = 0,07$). Par contre, si IDEAL avait utilisé le même critère combiné principal que dans TNT (incluant les AVC) ou dans PROVE-IT (incluant les procédures de revascularisation), la différence entre les deux groupes aurait été significative, avec un HR de 0,87 ($p = 0,02$) et un HR de 0,84 ($p < 0,001$), respectivement (12, 25).

Un autre aspect intéressant à discuter est celui de l'interruption prématurée des études pour des raisons éthiques et de son influence sur la signification statistique en ce qui concerne les critères de jugement des études (26). Deux cas récents sont exemplatifs à cet égard, celui de l'étude TNT (6, 11) et celui de l'étude ASCOT

(9, 14). Dans TNT, le comité de surveillance a décidé d'interrompre l'étude avant son terme en raison d'une différence hautement significative entre les deux groupes (atorvastatine 80 mg *versus* atorvastatine 10 mg) pour ce qui concerne le critère de jugement primaire (critère combinant les décès coronariens, les infarctus non mortels, les arrêts cardiaques avec ressuscitation et les AVC). Il en résulte qu'un critère secondaire important, celui de la mortalité totale, n'est pas significativement différent entre les deux groupes, ce qui a soulevé, de la part de certains, quelques réserves vis-à-vis du traitement à forte dose par atorvastatine (6, 11). A l'inverse, dans l'étude ASCOT, la surveillance des résultats par un comité indépendant a conduit à l'arrêt prématuré de l'étude en raison d'une différence hautement significative de mortalité entre les deux groupes, rendant non éthique la poursuite de l'essai en aveugle puisque cela aurait conduit à priver un groupe de patients (en l'occurrence la moitié d'entre eux) des effets bénéfiques potentiels d'un médicament actif. Cependant, la conséquence de cet arrêt prématuré de l'étude a été que le critère de jugement principal de l'essai, à savoir une combinaison des décès coronariens et des infarctus non mortels, y compris des infarctus silencieux, n'est pas sorti de façon significative pour démontrer la supériorité de l'approche amlodipine (+ périndopril) sur l'approche aténolol (+ bendrofluméthiazide) chez le sujet avec hypertension artérielle (9, 14). Nous avons déjà eu l'occasion de discuter précédemment les divers problèmes posés par ce compromis entre éthique et statistique (26).

CRITÈRES DE JUGEMENT SECONDAIRE

Les critères de jugement secondaire comportent soit des composantes individuelles du critère de jugement principal (dont souvent la mortalité cardio-vasculaire et/ou la mortalité totale), soit d'autres critères de jugement composites, plus ou moins complexes, soit encore d'autres critères (par exemple, testant la sécurité plutôt que l'efficacité). Selon la stratégie statistique considérée, l'analyse des critères de jugement secondaire ne peut être envisagée qu'à la condition exclusive que le critère de jugement principal atteigne le seuil de la signification statistique ou elle peut être réalisée même dans le cas où le critère de jugement principal n'est pas discriminant. Il s'agit là d'une différence essentielle et la stratégie d'analyse doit, théoriquement, être spécifiée d'emblée dans le plan statistique au départ de l'étude. Force est cependant de reconnaître que la démarche n'est pas toujours parfaitement transparente à la lecture

des articles originaux. Pourtant, les conclusions des études peuvent être complètement différentes selon le critère pris en considération. Ainsi, les études IDEAL (simvastatine 20-40 mg *versus* atorvastatine 80 mg) (7, 12), FIELD (fénofibrate *versus* placebo) (8, 13), ASCOT (amlodipine-perindopril *versus* aténolol-bendrofluméthiazide) (9, 14) et PROactive (pioglitazone *versus* placebo) (10, 15) ont utilisé des critères combinés d'évaluation secondaire qui se sont révélés statistiquement significatifs (études «positives») alors que les critères combinés d'évaluation primaire, qui avaient été privilégiés lors de la conception de l'essai, n'ont pas atteint le seuil de la signification statistique (études «négatives»).

CRITÈRES DE JUGEMENT COMPOSITES

La démonstration d'une réduction de la mortalité totale sous l'effet d'un médicament exige plusieurs conditions simultanées, parmi lesquelles une contribution majeure de la pathologie sur laquelle est sensé agir le médicament testé dans la mortalité globale et un nombre de décès escomptés suffisamment élevé dans le décours de l'étude, ce qui nécessite l'inclusion de patients à haut risque de décès, le recrutement d'un nombre suffisamment élevé de personnes et le suivi pendant une période assez longue. C'est la raison pour laquelle la mortalité globale est rarement prise comme critère de jugement principal dans les études de prévention cardiovasculaire (et jamais dans les différentes études publiées en 2005 et analysées ci-dessus) (Tableau I). C'était, par contre, le cas dans l'étude CHARM, tout au moins dans sa partie «Overall programme», qui avait testé l'effet du candesartan *versus* un placebo chez 7.601 patients décompensés cardiaques suivis pendant 37,7 mois (27, 28). Cette étude remplissait, en effet, les critères sus-mentionnés. Par contre, dans les trois sous-composantes du programme général, à savoir CHARM-Preserved, CHARM-Added et CHARM-Alternative, incluant chacune un nombre moins élevé de patients (et conduisant donc, par définition, à un nombre plus faible de décès-événements), le critère de jugement principal finalement choisi a été un critère composite combinant les décès cardiovasculaires et les hospitalisations pour une décompensation cardiaque (27, 28).

La mise en évidence d'un effet thérapeutique sur un critère d'évaluation isolé spécifique n'a guère d'intérêt, sauf si cet effet est associé à une réduction de la mortalité totale (29). Aussi, les grandes études de prévention cardio-vasculaire

font généralement appel à des critères de jugement composites. Ce type de critère offre certains avantages, ce qui explique leur utilisation de plus en plus large, mais comporte également certaines limites, voire expose à certains biais potentiels (17, 29).

AVANTAGES D'UN CRITÈRE DE JUGEMENT COMPOSITE

L'avantage principal de recourir à un critère de jugement composite est l'augmentation du nombre d'événements cliniques, ce qui permet de réduire le nombre de patients à inclure et de limiter la durée de l'étude. Ainsi, l'étude PROactive, la plus courte (2,8 années) et celle ayant inclus le moins de patients (quoique le nombre était loin d'être négligeable puisque 5.238 patients diabétiques de type 2 ont été randomisés) parmi les autres études analysées dans cet article est l'essai qui a également inclus la plus grande variété d'événements dans son critère de jugement principal, de façon à avoir une puissance statistique suffisante (10, 15) (Tableau I). Cette stratégie du choix du critère combiné principal permet donc d'augmenter la faisabilité des études et d'en réduire les coûts. De plus, des conclusions de type «evidence-base medicine» (EBM) peuvent être déduites plus rapidement, ce qui offre l'avantage d'orienter favorablement les choix thérapeutiques dès que les preuves sont disponibles et, donc, d'en faire bénéficier, aussitôt que possible, les patients potentiellement concernés. Ces avantages ne doivent cependant pas occulter les désavantages éventuels (17, 29).

LIMITES D'UN CRITÈRE DE JUGEMENT COMPOSITE

Le recours à un critère de jugement composite a comme désavantage l'hétérogénéité des composantes faisant partie du critère combiné, notamment en terme de signification clinique, comme déjà mentionné (17). Ainsi, prévenir un décès a un impact supérieur par rapport à éviter la survenue d'un angor ou le recours à une angioplastie. De plus, il peut arriver que le médicament étudié ait un effet attendu variable sur chacune des composantes, ce qui peut aboutir à une dilution possible de l'effet aboutissant à une différence finale qui n'est pas significative sur le plan statistique. Par exemple, dans l'étude PROactive, le groupe pioglitazone présente moins d'événements coronariens ou cérébro-vasculaires que le groupe placebo, alors qu'il n'y a pas de différence en ce qui concerne les manœuvres de revascularisation (y compris des membres inférieurs) et que la tendance est plutôt dans le sens opposé (incidence plus élevée dans le groupe pioglitazone, non significative cependant) pour ce qui concerne les amputations des

membres inférieurs. Ceci explique, au moins en partie, pourquoi le critère de jugement principal (qui prenait en compte ce paramètre) n'est pas significatif alors que le critère de jugement secondaire principal (qui ne prenait pas en compte ce paramètre) atteint le seuil de la signification statistique (9, 11). Or, il est bien connu que l'insuffisance artérielle ne joue qu'un rôle partiel dans la problématique du pied diabétique où interviennent aussi la composante neuropathique et surtout la composante infectieuse, par définition non influencées par le médicament testé, en l'occurrence la pioglitazone.

Le choix des composantes incluses dans le critère combiné peut également conduire à un biais potentiel. Ainsi, ce choix, souvent guidé par les intérêts de l'industrie pharmaceutique, paraît clairement influencé par le type de médicament évalué. Ainsi, dans l'étude VALUE comparant le valsartan et l'amlodipine chez des sujets hypertendus (20), la décompensation cardiaque a été incluse par les promoteurs dans le critère combiné principal, choix qui inévitablement permettait de donner un certain avantage au valsartan par rapport à l'amlodipine au vu du mécanisme d'action particulier déjà rappelé pour les deux molécules. En l'occurrence, cependant, cet avantage, s'il a bien été confirmé sur ce point spécifique, n'a pas permis d'influencer significativement le critère d'évaluation principal en faveur du valsartan par comparaison à l'amlodipine (20). Au contraire, dans l'étude PROactive, l'insuffisance cardiaque n'a pas été prise en considération dans le critère combiné principal, choix qui aurait risqué de pénaliser d'emblée la pioglitazone par rapport au placebo puisque les glitazones sont connues pour favoriser la survenue d'oedèmes et de précipiter une décompensation cardiaque chez des sujets à risque (10, 15). Si les cas de décompensation cardiaque avaient été inclus dans le critère de jugement principal, la tendance favorable observée sous pioglitazone aurait été annulée, ainsi que discuté dans l'éditorial accompagnant l'article original (22).

CONCLUSION

Le choix des critères d'évaluation dans les essais cliniques est devenu crucial, aussi bien pour l'industrie pharmaceutique que pour les commissions d'évaluation des médicaments. Ceci est d'autant plus important que la survenue d'événements cliniques est de plus en plus basse (compte tenu d'une meilleure prise en charge des différents facteurs de risque dans la population générale), que les essais s'adressent à des

patients avec un risque global plus faible (puisque'il ne peut plus être considéré éthique de faire des essais, *a fortiori versus* placebo, chez des patients à haut risque) et que la différence entre deux produits est de plus en plus marginale (en particulier lorsqu'il s'agit de comparer deux molécules de la même classe). L'analyse de la littérature récente dans le domaine de la prévention cardio-vasculaire démontre, à suffisance, la disparité des critères utilisés entre les études, met en évidence des discordances dans les résultats entre les critères d'évaluation primaires et secondaires et révèle même, dans certains cas, des modifications des critères de jugement principaux en cours d'étude.

Le clinicien confronté aux résultats des études cliniques publiées doit être suffisamment au fait de cette problématique pour pouvoir analyser les résultats de façon critique, qu'ils soient positifs ou négatifs, en toute connaissance de cause. Quoi qu'il en soit, le temps paraît révolu où l'analyse des critères de jugement secondaire n'était pas acceptée en l'absence de différence statistiquement significative en ce qui concerne le critère de jugement principal. C'est certainement le cas dans le domaine de la prévention cardio-vasculaire où les différents critères composites utilisés, primaires ou secondaires, sont généralement très proches l'un de l'autre et tous porteurs d'une signification clinique indéniable.

BIBLIOGRAPHIE

1. Scheen AJ.— Evidence-based medicine. Apport des essais cliniques contrôlés. *Rev Med Liège*, 2000, **55**, 216-219.
2. Delvenne C, Pasleau F.— Comment résoudre en pratique un problème diagnostique ou thérapeutique en suivant une démarche EBM ? *Rev Med Liège*, 2000, **55**, 226-232.
3. Van den Bruel A, Chevalier P, Vermeire E, et al.— Evidence-Based Medicine. Otite moyenne chez l'enfant. Comment formuler une question PICO ? *Rev Med Liège*, 2004, **59**, 671-675.
4. Scheen AJ.— La mesure de l'efficacité thérapeutique, étape essentielle dans la médecine factuelle. *Rev Med Liège*, 2000, **55**, 206-210.
5. Kuller LH.— Clinical trials : surrogate endpoints or hard endpoints ? *Am J Cardiol*, 2001, **88** (suppl), 59E-61E.
6. LaRosa JC, Grundy SM, Waters DD, et al.— Intensive lipid lowering with atorvastatin in patients with stable coronary disease. *N Engl J Med*, 2005, **352**, 1425-1435.
7. Pedersen TR, Faergeman O, Kastelein JJP, et al. The Increment Decrease in End Points Through Aggressive Lipid Lowering (IDEAL) Study Group.— High-dose atorvastatin vs usual-dose simvastatin for secondary prevention after myocardial infarction. The IDEAL study : a randomized controlled trial. *JAMA*, 2005, **294**, 2437-2445.

8. The FIELD study investigators.— Effects of long-term fenofibrate therapy on cardiovascular events in 9795 people with type 2 diabetes mellitus (the FIELD study) : randomised controlled trial. *Lancet*, 2005, **366**, 1849-1861.
9. Dahlöf G, Sever P, Poulter N, et al.— Prevention of cardiovascular events with an antihypertensive regimen of amlodipine adding perindopril as required versus atenolol adding bendroflumethiazide as required, in the Anglo-Scandinavian Cardiac Outcomes trials : Blood Pressure Lowering Arm (ASCOT-BPLA) : a multicentre randomised controlled trial. *Lancet*, 2005, **366**, 895-906.
10. Dormandy J, Charbonnel B, Eckland DJA, et al, on behalf of the PROactive investigators.— Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitazone Clinical Trial In macroVascular Events) : a randomised controlled trial. *Lancet*, 2005, **366**, 1279-1289.
11. Scheen AJ.— Treating to New Targets : plaidoyer pour une valeur cible de cholestérol LDL égale ou inférieure à 75 mg/dl chez tout patient coronarien. *Rev Med Liège*, 2005, **60**, 264-267.
12. Scheen AJ.— L'étude IDEAL comparant simvastatine 20-40 mg versus atorvastatine 80 mg en prévention après un infarctus du myocarde : entre deux idées de l'idéal. *Rev Med Liège*, 2006, **61**, 53-59.
13. Radermecker RP, Scheen AJ.— L'étude FIELD de protection cardio-vasculaire avec le fénofibrate chez le patient diabétique de type 2. *Rev Med Liège*, 2005, **60**, 957-961.
14. Krzesinski JM.— Première démonstration d'une supériorité des agents antihypertenseurs modernes par rapport aux molécules plus anciennes dans la prévention cardio-vasculaire : résultats du bras "abaissement de pression artérielle" de l'étude ASCOT. *Rev Med Liège*, 2005, **60**, 820-826.
15. Scheen AJ, Lefèbvre PJ.— L'étude PROactive : prévention secondaire des accidents cardio-vasculaires par la pioglitazone chez le patient diabétique de type 2. *Rev Med Liège*, 2005, **60**, 896-901.
16. Freemantle N.— Interpreting the results of secondary end points and subgroup analyses in clinical trials : should we lock the crazy aunt in the attic ? *BMJ*, 2001, **322**, 989-991.
17. Freemantle N, Calvert M, Wood J, Eastaugh J, Griffin C.— Composite outcomes in randomized trials. Greater precision but with greater uncertainty ? *JAMA*, 2003, **289**, 2554-2559.
18. Nissen SE, Nicholls SJ, Sipahi I, et al.— Effect of very high-intensity statin therapy on regression of coronary atherosclerosis. The ASTEROID trial. *JAMA*, 2006, **295**, online March 13 (doi:10.1001/jama.295.13.jpc60002).
19. Taylor A, Shaw LJ, Fayad Z, et al.— Tracking atherosclerosis regression : a clinical tool in preventive cardiology. *Atherosclerosis*, 2005, **180**, 1-10.
20. Julius S, Kjeldsen SE, Weber M for the VALUE trial group.— Outcomes in hypertensive patients at high cardiovascular risk treated with regimens based on valsartan or amlodipine : the VALUE randomised trial. *Lancet*, 2004, **363**, 2022-2031.
21. The FIELD Study Investigators.— The need for a large-scale trial of fibrate therapy in diabetes : the rationale and design of the Fenofibrate Intervention and Event Lowering in Diabetes (FIELD) study. *Cardiovascular Diabetology*, 2004, **3**, 9 (doi:10.1186/1475-2840-3-9).
22. Yki-Järvinen H.— The PROactive study : some answers, many questions (Editorial). *Lancet*, 2005, **366**, 1241-1242.
23. Freemantle N.— How well does the evidence on pioglitazone back up researchers' claims for a reduction in macrovascular events ? *BMJ*, 2005, **331**, 836-838.
24. Cannon CP, Braunwald E, McCabe CH, et al.— Comparison of intensive and moderate lipid lowering with statins after acute coronary syndromes. *N Engl J Med*, 2004, **350**, 1495-1504.
25. Cannon CP.— The IDEAL cholesterol. Lower is better (Editorial). *JAMA*, 2005, **294**, 2492-2494.
26. Scheen AJ.— Interruption prématurée des études ASCOT et CARDS de prévention cardio-vasculaire avec l'atorvastatine chez le sujet hypertendu ou diabétique : compromis entre éthique et statistique en médecine factuelle. *Rev Med Liège*, 2003, **58**, 585-590.
27. Pfeffer MA, Swedberg K, Granger CB, et al.— Effects of candesartan on mortality and morbidity in patients with chronic heart failure : the CHARM-Overall programme. *Lancet*, 2003, **362**, 759-766.
28. Kulbertus H.— L'étude CHARM. *Rev Med Liège*, 2003, **58**, 646-652.
29. Lubsen J, Kirwan BA.— Combined endpoints : can we use them ? *Stat Med*, 2002, **21**, 2959-2970.

Les demandes de tirés à part sont à adresser au
Pr A.J. Scheen, Département de Médecine, CHU Sart
Tilman, 4000 Liège, Belgique